基于轻量级MobileNet V3-YOLOv4的生长期菠萝成熟度分析

李阳德[1,2]，马晓慧[3*]，王　骥[3,4*]

（1. 广东省电信规划设计院有限公司第四分公司，广东湛江 524088；2. 华中科技大学 电子信息与通信学院，湖北武汉 430070；3. 广东海洋大学电子与信息工程学院，广东湛江 524088；4. 广东省智慧海洋传感网及其装备工程技术研究中心，广东湛江 524088；）

摘　要：［**目的/意义**］菠萝的贮藏性与成熟度相关，菠萝采摘前对其成熟度进行识别尤为重要。本研究目的在于提出一种新型网络模型，提高菠萝成熟度自动识别的准确率和速度。［**方法**］首先针对菠萝训练数据集样本少与实时性差等不足，利用在自然环境下拍摄的菠萝照片，自建了种植区场景菠萝成熟度分析数据集。之后将YOLOv4骨干网络替换成轻量级网络MobileNet V3，提出了轻量级的MobileNet V3-YOLOv4网络。同时训练了原YOLOv4模型、MobileNet V1-YOLOv4模型、MobileNet V2-YOLOv4模型以及Faster R-CNN、YOLOv3、SSD300、Retinanet、Centernet等五种不同的单、双阶段网络模型，并对比模型的评价指标，分析本文模型的优越性。［**结果和讨论**］试验结果表明，MobileNet V3-YOLOv4训练时间为11,924 s，参数量为53.7 MB，训练好的MobileNet V3-YOLOv4在验证集的平均精度均值（mean Average Precision，mAP）为90.92%，对于黄熟期菠萝和青熟期菠萝两种类别的检测精确率（Precision）分别为100%和98.85%，平均精度（Average Precision，AP）值分别为87.62%、94.21%，召回率（Recall）分别为77.55%、86.00%，F_1分数（F_1 Score）分别为0.87和0.92，推理速度（Frames Per Second，FPS）80.85 img/s。［**结论**］本研究提出的MobileNet V3-YOLOv4实现了在降低训练速度、减小参数量的同时，提高了菠萝成熟度识别的精度和推理速度，满足实际检测需求。

关键词：菠萝成熟度；骨干网络；MobileNet V3-YOLOv4；Faster R-CNN；SSD300；Retinanet；Centernet；轻量级

中图分类号：TP391　　　　**文献标志码：**A　　　　**文章编号：**SA202211007

引用格式：李阳德, 马晓慧, 王骥. 基于轻量级MobileNet V3-YOLOv4的生长期菠萝成熟度分析[J]. 智慧农业(中英文), 2023, 5(2): 35-44.

LI Yangde, MA Xiaohui, WANG Ji. Pineapple maturity analysis in natural environment based on Mobilenet V3-YOLOv4[J]. Smart Agriculture, 2023, 5(2): 35-44.

1　引　言

菠萝是一种热带水果，在中国的主要产地为海南、广东等南方省份，种植面积大、产量多。若菠萝采摘作业较晚，极有可能导致菠萝腐烂，而采摘过早则会影响到食用状态和口感[1]。菠萝从采摘地运往全国各地，因运输目的地远近不同，运输周期与储藏周期也不同。菠萝成熟度越高，菠萝的贮藏性越差，运输周期相应要求也越短，故需要在采摘前对菠萝的成熟度做出准确的判断，使得菠萝能在合适的成熟度进行采摘与运输到不同距离的目的地。

传统的水果成熟度分析方法一般是通过人工判别、对果实成分进行分析，或者通过检测挥发的气体成分与浓度进行识别[2]。传统的方法虽然对果实的成熟程度分析比较准确，但会对果实的成长进行干预，且过程复杂、成本高、效率低，不具备实时

收稿日期：2022-11-17

基金项目：广东省人工智能领域专项（2019KZDZX1046）

作者简介：李阳德，高级工程师，研究方向为物联网应用及智慧城市建设研究。E-mail：liyangde@gpdi.com

*通信作者：1. 马晓慧，硕士，研究方向为农业人工智能。E-mail：2068813294@qq.com；

2. 王　骥，硕士，教授，研究方向为无线传感器网络、海洋物联网与人工智能等；E-mail：13902576499@163.com

性[3]。随着计算机视觉技术的不断深入发展，使用卷积神经网络（Convolutional Neural Network，CNN）的图像识别算法在农业中应用逐渐成熟，成为解决视觉任务的有效模型[4]。许多学者将CNN的图像识别算法应用于水果的识别检测领域，通过提取图片或者视频中果实的颜色、纹理、形状特征，实现果实的识别分析。Gai等[5]提出了一种适合检测体积较小的樱桃水果改进的YOLOv4深度学习算法来检测樱桃果实，改进的YOLOv4模型网络得到的平均精度均值（mean Average Precision，mAP）值比普通的YOLOv4高0.15。Chen等[6]利用图像的大小、高度、宽度等特征，通过YOLOv3网络对指定的水果进行连续跟踪，并在跟踪过程中检测水果的质量，对6000幅水果图像进行了测试，准确率高达88%。Kuznetsova等[7]提出了预处理和后处理技术使YOLOv3算法应用于苹果收获机器人机器视觉系统，该系统对苹果平均检测时间为19 ms，其中被误认为苹果的物体比例为7.8%，未被识别的苹果比例为9.2%。Zhang等[8]设计了一个13层的CNN，对梨、苹果以及腐烂水果进行分类识别，准确率高达94.94%。Chaikaew等[9]研究了一种应用神经网络的菠萝分选机，该分选机基于菠萝的颜色，将菠萝的成熟度分为未熟菠萝、部分熟菠萝和完全熟菠萝，其对未成熟菠萝的识别准确率为79%，对部分成熟菠萝的识别准确率为82%，对完全成熟菠萝的准确率为100%。Cuong等[10]为了实现菠萝成熟度的预测，提出了一种改进的Tiny YOLOv4模型，模型的识别准确率高达98.26%。张星等[11]为实现果实拾捡机器人在各种因素影响下的田间复杂环境对单类别菠萝的快速准确识别，提出采用深度学习的深层残差网络改进YOLOv3卷积神经网络结构，该模型的识别准确率达95%左右。Liu等[12]为实现菠萝果实的智能采摘，提出了一种基于双目立体视觉和改进的YOLOv3模型的自然环境中菠萝检测与定位方法，改进后的网络在轻微遮挡的测试集上F_1分数（F_1 Score）和平均精度（Average Precision，AP）值分别为93.18%和97.55%，在闭塞程度的加重测试集上，F_1 Score和AP值分别下降到了89.15%和91.47%。

综上，目前目标检测算法研究状态，具有快速、准确、无损、高工作效率等特点，文献［5］—［10］中的识别方法的数据集图像均为背景单一的图片，适用于实验室等无遮挡的理想场景，而生长状态的菠萝真实场景的现场检测却存在枝叶遮挡、光照不均、阴影重叠等问题，文献［11］—［12］中虽然实现了复杂场景下菠萝果实的识别，但只是对单类别的菠萝识别，不具备通用性，实际检测工作中为减少成本，要求网络处理速度快、吞吐量大、识别精度高、通用性强，一般卷积神经网络很难达到指标要求。针对上述问题，本研究以YOLOv4[13]网络为基础，结合MobileNet V3提出了新型优化网络，即MobileNet V3-YOLOv4模型，该模型具有较少训练时间和识别准确率高的优点，满足真实场景下对菠萝成熟度快速识别分析需求。

2 材料与方法

2.1 数据集构建

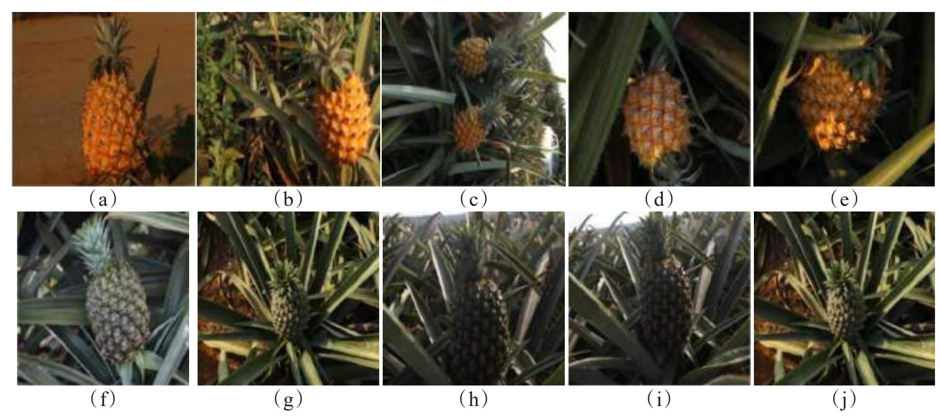
2.1.1 数据获取

网络训练数据采集时间为2023年3月25日，采集地点为湛江市徐闻县曲界镇菠萝的海，采集设备为SONY DSC-RX100M7，采集图片的像素为720×480。根据菠萝的颜色、形状、纹理等特征，本研究将自然环境需要采摘的菠萝划分为青熟期和黄熟期。对采集的图片进行初步筛选，将过度曝光、过度模糊、不含目标的照片舍弃。经过筛选后得到1580幅图片，其中黄熟期菠萝图片674幅、青熟期菠萝图片906幅，随机选择1264幅图片作为训练集、158幅作为验证集、158幅作为测试集。

数据样本的采集发生在环境复杂的菠萝种植田中，且在采集数据时，拍摄的角度、时间、光照强度各有不同。采集的样本包含下列情况：无遮挡无阴影光照适中的合适样本857幅，枝叶遮挡的样本203幅，光照不均形成阴影的样本232幅，光照强度较强或较弱影响下的样本112幅，多种影响因素交叉共存下的样本176幅。自然环境下采集的菠萝图片，保证了样本的多样性。图1展示了受到各类因素影响的各类型样本具体情况。

2.1.2 数据标注

为了满足网络的训练要求，使用图像标注软件



注：(a)(f)为无遮挡无阴影光照适中的合适样本；(b)(g)为光照不均的样本；(c)(h)为枝叶遮挡的样本；(d)(i)为阴影遮挡的样本；(e)(j)为枝叶遮挡的样本、光照不均且阴影覆盖的样本

图1 不同采集环境下的菠萝样本图

Fig. 1 Pineapple samples in different collection environments

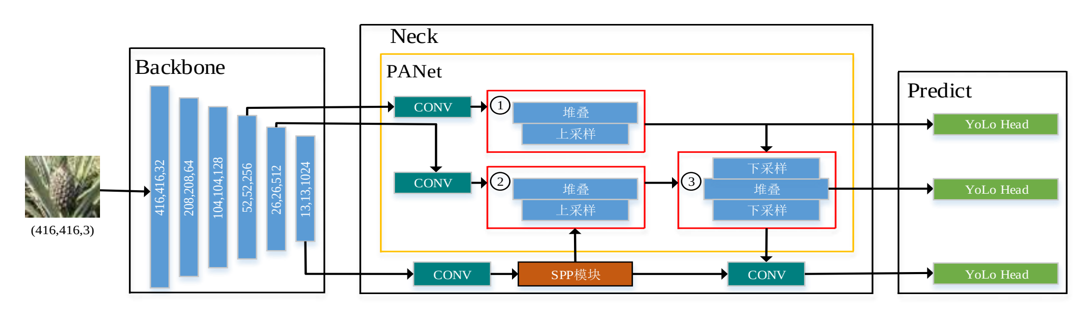
labelimg对图像进行标注，生成*.xml标签文件，对于图像中目标被遮挡的面积超过70%、目标表面的阴影过重无法分辨本身色彩、形状较小距离较远无法分辨边缘轮廓的图像不予标注。其中青熟期菠萝、黄熟期菠萝对应的标签名分别为Growing pineapple和Mature pineapple。

2.2 MobileNet V3-YOLOv4网络构建

YOLOv4网络模型存在结构复杂、参数量庞大等问题，导致训练时间长、识别误差大等，而在农田场景下，一般架设极少数的监控设备实现农田监测，故要求监控设备的视觉算法，在使用极少算力的条件下实现菠萝的快速检测。基于此，本研究引入了MobileNet V3模型替代YOLOv4原本的骨干特征网络提出了MobileNet V3-YOLOv4。

2.2.1 YOLOv4和MobileNet

YOLOv4网络结构主要包括三个部分：①CSPDarknet53[14]主干特征提取网络，可实现图片特征的提取，获得三个有效的初步特征层，三个初步有效特征层的大小（高，宽，通道数）分别用于检测小物体的（52，52，256），检测中等物体的（26，26，512），检测大物体的（13，13，1024）；②颈部结构（neck）由空间金字塔池化结构（Spatial Pyramid Pooling，SPP）模块与PAnet（Path Aggregation Network，路径聚合网络）两部分共同组成，SPP模块可以提升感受野，PAnet则是由YOLOv3[15]的特征金字塔（Feature Pyramid Networks，FPN）[16]改进而成，该结构的功能是进行加强的特征提取，利用加强特征提取网络，可以对三个初步的有效特征层进行特征融合，提取出更好的特征，获得三个更有效的有效特征层；③预测网络，由三个YOLO Head模块组成，该结构可利用颈部结构获取到的更有效的有效特征层进行回归和分类，获得最终的预测结果。YOLOv4的结构如图2所示。



注：图中SPP模块用4个大小分别为13×13、9×9、5×5和1×1的卷积核将特征层的输出进行最大池化操作；①、②、③为PANet经过五次卷积和特征融合获得的特征层，三大特征层对应的大小分别为52×52、26×26和13×13，三个特征层将被送到Head模块用于预测

图2 YOLOv4网络结构

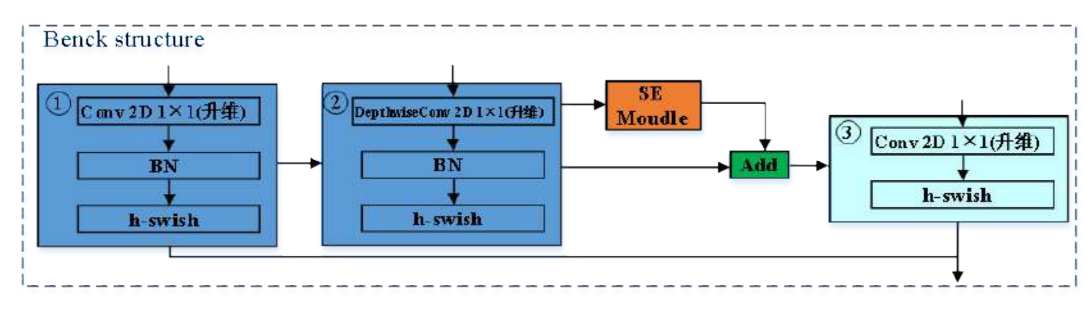
Fig.2 Structure of YOLOv4

MobileNet是谷歌公司2017年提出的一种专注于移动端或者嵌入式设备的CNN网络，其特点是轻量、快速[17]。该网络首次用深度可分离卷积模块替代了普通卷积的方式。MobileNet V3则是结合深度可分离模块和残差结构提出了一种benck结构[18]，其结构如图3所示。Benck结构的主干部分，

首先利用1×1卷积进行升维，然后利用3×3深度可分离卷积进行特征提取，再利用1×1卷积降维，残差部分则是输入和输出直接相连。Benck结构中用h-swish激活函数替代了RELU激活函数，h-swish函数的表示为公式（1），并在可分离卷积和降维卷积之间增加了SE（Squeeze Excitation）模块，其结构如图4所示，由一个全局池化结构（Global Pooling）、两个全连接层（Fully Connected Layers，FC）、一个RELU激活函数和一个h-sigmoid函数组成，h-sigmoid函数替代了sigmoid函数，具体表示为公式（2）。

$$h\text{-}swish[x] = x\frac{RELU6(x+3)}{6} \tag{1}$$

$$h\text{-}sigmoid[x] = \frac{RELU6(x+3)}{6} \tag{2}$$



注：①为普通卷积用于升维；②为深度卷积；③为普通卷积用于降维；结构③不需要BN块进行归一化处理

图3 Benck结构

Fig. 3 The structure of the Benck

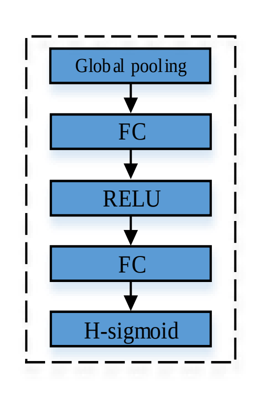


图4 SE模块

Fig.4 SE module

2.2.2 MobileNet V3-YOLOv4主干特征网络构建

MobileNet V3-YOLOv4使用benck结构替代YOLOv4骨干网络中CSPDarknet的Resblock，改进后的主干网络结构如图5所示。当输入的图片为（416，416，3）时，先经过一个大小3×3，步长为2的卷积提取出（280，208，16）的特征层，再经过多次的benck，获得三个有效的初步特征层，分别位于网络的中间层、中下层和底层，其对应特征层大小分别为（52，52，40）、（26，26，112）和（13，13，160），用于输入下层的加强特征提取网络，进行更有效的特征提取。

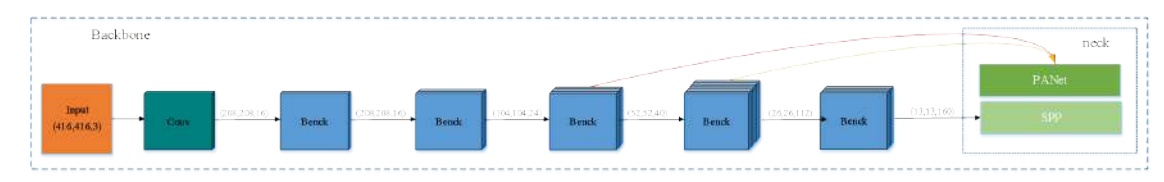


图5 MobileNet V3-YOLOv4的主干网络结构

Fig. 5 The backbone network structure of MobileNet V3-YOLOv4

此外，本研究还构建了MobileNet V1-YOLOv4和MobileNet V2-YOLOv4两种网络结构作为对比结构。

2.2.3 目标检测评价指标

本研究的标签分为Growing pineapple和Mature pineapple 2类，以AP、召回率（Recall）、F_1 Score、精确率（Precision）和mAP值作为目标检测评价指标，相应计算方法如公式（3）～（6）所示。

$$Precision = \frac{TP}{TP + FN} \tag{3}$$

$$Recall = \frac{TP}{TP + FP} \tag{4}$$

$$F_1\,Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

$$\mathrm{mAP} = \frac{1}{\mathrm{C}}\sum_{i \in C} AP_i \qquad (6)$$

其中，TP为目标被分为正类且分类正确的数量，个；TN为目标被分为负类且分类正确的数量，个；FP为目标被分为正类但分类错误的数量，个；C为分类的类别数，本研究C为2。

2.3 菠萝成熟度检测算法识别工作流程

本研究首先通过使用摄像机人工拍摄的方式完成图片的采集工作，同时对图片进行初步筛选，将不符合网络训练的图片舍弃；接着按照菠萝的成熟指标（本研究主要依据颜色）进行判别，对不同成熟度的菠萝照片进行标注，标注的标签分为黄熟度和青熟度；将标注完成的图片作为数据集，输入网络进行训练，图片先经过网络的主干结构，获取到初步特征，然后经过网络的颈部结构，对图片的初步特征进行进一步处理，获取到进一步的加强特征，将加强特征输入至头部结构，经过回归和分类，获得预测结果，生成成熟的网络模型；再将成熟的网络模型输入到评价模型，对网络模型进行评价指标的计算，最后输出网络的评价指标。菠萝成熟度检测算法的工作流程如图6所示。

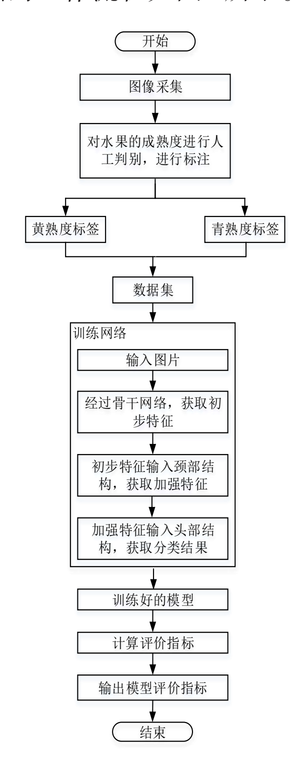


图6 菠萝成熟度检测算法工作流程

Fig.6 Flowchart of pineapple ripeness test algorithm

3 试验结果与分析

本研究在YOLOv4网络基础上进行了改进，构建新的网络模型MobileNet V3-YOLOv4。为验证改进方法对菠萝成熟度检测的有效性，进行了模型性能对比试验、模型分类训练对比试验以及复杂环境下菠萝检测准确率试验，记录了不同模型训练自建菠萝数据集训练集的性能指标和评价指标。性能指标包括训练时间和模型参数，是评估模型复杂度的重要指标，通过对比不同模型的性能指标，验证改进后的方法是否比原网络降低了模型的复杂度；模型训练评价指标包括*Precision*值、AP值、推理速度（Frames Per Second，FPS）等，是评估模型精度和推理速度的重要指标，通过对比不同模型的评价指标，验证改进后的方法是否优化了原模型的精度和推理速度。同时为了验证改进后的模型在复杂环境下对菠萝果实的检测效果，用训练好的模型对菠萝数据集的测试集图片进行了随机检测。

3.1 试验环境与参数设置

试验的软件环境为：ubuntu 16.04，Anaconda3，pytorch1.2.0，CUDA10.0，cudnn7.6.4。

硬件环境为：GPU：NVIDA GeForce GTX 2080ti，CPU：Inter（R）Core（TM）i5-8400 CPU @2.80 GHz；编译语言为Python3.7。网络训练前设置的初始参数如表1所示。

表1 菠萝成熟度试验网络初始参数设置

Table 1 Setting of network initial parameter

初始参数	初始参数值
训练周期(epoch)/轮	300
批处理	8
初始学习率	0.01
动量参数	0.937

3.2 模型性能对比

为评估模型复杂度，选用Faster R-CNN、YOLOv3、SSD300、Retinanet、Centernet、YOLOv4、MobileNet V1-YOLOv4、MobileNet V2-YOLOv4和本研究构建的MobileNet V3-YOLOv4九种模型参与模型性能对比分析，性能指标包括训练时间、饱和周期和模型参数。结果显示，本研究构建的MobileNet V3-YOLOv4模型的训练时间为11,924 s，平

均每轮的训练时间为39.75 s，在第256轮达到饱和，参数量为53.7 MB，达到饱和的时间则比YOLOv4减少了3500.74 s，参数量只占原网络的22%。九种模型性能指标对比结果如表2所示。

表2 菠萝成熟度试验各模型性能指标对比

Table 2 Comparison of performance indexes of pineapple maturity test models

序号	网络	训练时间/s	平均每轮训练时间/s	饱和周期/轮	参数量/MB
1	Faster R-CNN	31,881	106.27	180	108.0
2	YOLOv3	20,820	69.40	170	235.0
3	SSD300	15,728	52.42	100	91.1
4	Retinanet	10,128	33.76	42	138.0
5	Centernet	16,462	54.87	160	124.0
6	YOLOv4	19,172	63.91	214	244.0
7	MobileNet V1-YOLOv4	11,000	36.67	195	51.1
8	MobileNet V2-YOLOv4	12,068	40.23	227	46.5
9	**MobileNet V3-YOLOv4**	**11,924**	**39.75**	**256**	**53.7**

3.3 模型训练结果对比对比

3.3.1 训练结果分析

为了验证MobileNet V3-YOLOv4模型的分类预测性能，利用 *Recall* 分数、F_1 Score、*Precision* 和AP值的四个指标对不同成熟度的菠萝进行分类识别。同时本研究训练了YOLOv4、MobileNet V1-YOLOv4、MobileNet V2-YOLOv4 以及 Faster R-CNN、YOLOv3、SSD300、Retinanet、Centernet 等不同的单、双阶段网络模型，将不同模型的训练结果进行对比分析，训练结果如表3所示。

表3 单类目标检测试验各模型评价指标结果对比

Table 3 Comparison of the results of each model evaluation index in single-class target detection test

序号	网络	青熟期菠萝				黄熟期菠萝			
		AP/%	F_1 Score	*Recall*/%	*Precision*/%	AP/%	F_1 Score	*Recall*/%	*Precision*/%
1	Faster R-CNN	72.36	0.76	78.09	73.33	94.41	0.94	94.95	93.62
2	YOLOv3	85.19	0.78	67.45	93.65	89.31	0.84	75.87	93.22
3	SSD300	68.00	0.81	80.37	80.70	94.20	0.96	95.83	95.64
4	Retinanet	84.65	0.8	71.96	89.24	89.10	0.84	75.14	96.35
5	Centernet	73.25	0.36	21.70	98.55	90.02	0.66	49.84	99.37
6	YOLOv4	89.72	0.85	85.00	95.51	81.56	0.74	61.00	93.75
7	MobileNet V1-YOLOv4	92.53	0.92	86.00	98.85	89.14	0.82	73.47	92.31
8	MobileNet V2-YOLOv4	92.73	0.89	89.00	98.89	87.61	0.80	79.59	100.00
9	**MobileNet V3-YOLOv4**	**94.21**	**0.92**	**86.00**	**98.85**	**87.62**	**0.87**	**77.55**	**100.00**

由表3可知，Faster R-CNN、YOLOv3、SSD300、Retinanet和Centernet五种不同的单、双阶段网络模型虽然对黄熟期菠萝的检测有较好的效果，但对青熟期菠萝的检测效果较差。五种单双网络中，YOLOv3对青熟期的菠萝检测效果最好，但AP值低于YOLOv4系列的所有网络。YOLOv4为基础的系列网络中，MobileNet V3-YOLOv4的精确率、AP值、F_1 Score明显高于YOLOv4网络，青熟期AP值提升了4.49%、F_1 Score提升了0.07、*Recall* 值提升了1%、精确率提升了3.34%；而黄熟期AP值提升了6.06%、F_1 Score 提升了0.13，*Recall* 值提升了16.55%，*Precision* 提升了6.25%。由于黄熟期菠萝颜色特征明显，易与背景信息区分，故改进后的MobileNet V3-YOLOv4网络对黄熟期菠萝的检测时，*Precision* 高达100.00%，比青熟期菠萝高出1.15%，但青熟期菠萝的 *Recall* 分数较黄熟期菠萝高出8.45%，青熟期菠萝的 F_1 Score较黄熟期菠萝高出5%，不同成熟度的菠萝果实在复杂环境下的AP值均在85%以上，但青熟期的AP值比黄熟期高出6.59%。

表3中的AP值、*Recall* 值和 *Precision* 只能验证网络对单类目标的检测效果，而评价网络的性能指标还有FPS值与mAP值，FPS表示网络的推理速度，即网络每秒识别的图片数量，mAP值为各类别AP值的平均值，可以验证网络对多类目标的检测效果，测试结果如表4所示。由表4可知，五种

单双阶段网络中mAP值最高的为YOLOv3，FPS只有51.33 img/s，而推理速度最快的SSD300 mAP值的表现最差。MobileNet V3-YOLOv4的mAP值比YOLOv4高出5.28%，比YOLOv3高出3.67%，推理速度比YOLOv4高出40.28 img/s，比SSD300高出8.91 img/s。

表4　多类目标检测试验各模型评价指标结果对比

Table 4 Comparison of evaluation index results of various models in multi-class target detection test

序号	网络	FPS/(img·s^{-1})	mAP
1	Faster R-CNN	15.22	83.39
2	Yolov3	51.33	87.25
3	SSD300	71.94	81.1
4	Retinanet	27.17	86.88
5	Centernet	68.07	81.64
6	YOLOv4	40.57	85.64
7	MobileNet V1-YOLOv4	69.97	90.83
8	MobileNet V2-YOLOv4	75.95	90.17
9	**MobileNet V3-YOLOv4**	**80.85**	**90.92**

3.3.2　讨论

通过分析表2~4可知，MobileNet V3-YOLOv4的各个方面都明显优于YOLOv4模型，虽然有些指标略低于MobileNet V1-YOLOv4和MobileNet V2-YOLOv4，但考虑到实际应用中会存在不同的限制条件，本研究从以下几个方面进行讨论分析：

（1）为了验证本方法对单类别菠萝目标的检测效果，将模型的AP值作为指标。MobileNet V3-YOLOv4对两种成熟度菠萝的AP值都明显高于YOLOv4系列的其他三种网络。不同成熟度的菠萝果实在大小和形状特征上相差较小，改进后的MobileNet V3-YOLOv4通过区分果实的颜色特征及有细微变化的纹理特征，实现了对单类别目标的判断。

（2）为了验证本方法对多类别目标的检测效果，将*Precision*和mAP值作为指标。由于黄熟期的菠萝颜色特征更明显，纹理特征相对较突出，黄熟期菠萝目标与背景的对比度比青熟期更大，对比的其他网络只能保证网络对单类别的黄熟期菠萝目标的检测效果，实际生产中则需要考虑的是多类别目标的检测效果。MobileNet V3-YOLOv4和MobileNet V2-YOLOv4的对黄熟期菠萝的*Precision*最高可达100%，但MobileNet V3-YOLOv4的mAP比MobileNet V2-YOLOv4提升了0.75%，兼顾了对两种成熟度菠萝的检测效果。

（3）从计算资源考虑和模型复杂度考虑，将模型的参数量和训练时间作为指标。MobileNet V3-YOLOv4的参数量比YOLOv4网络减少了37.80%，MobileNet V3-YOLOv4的训练时间比YOLOv4减少了7248 s，比MobileNet V2-YOLOv4减少了144 s，与MobileNet V1-YOLOv4相差较少，其*Precision*、AP等都明显增长，保证了网络在保持精度的同时，降低了模型的复杂度和运算成本。

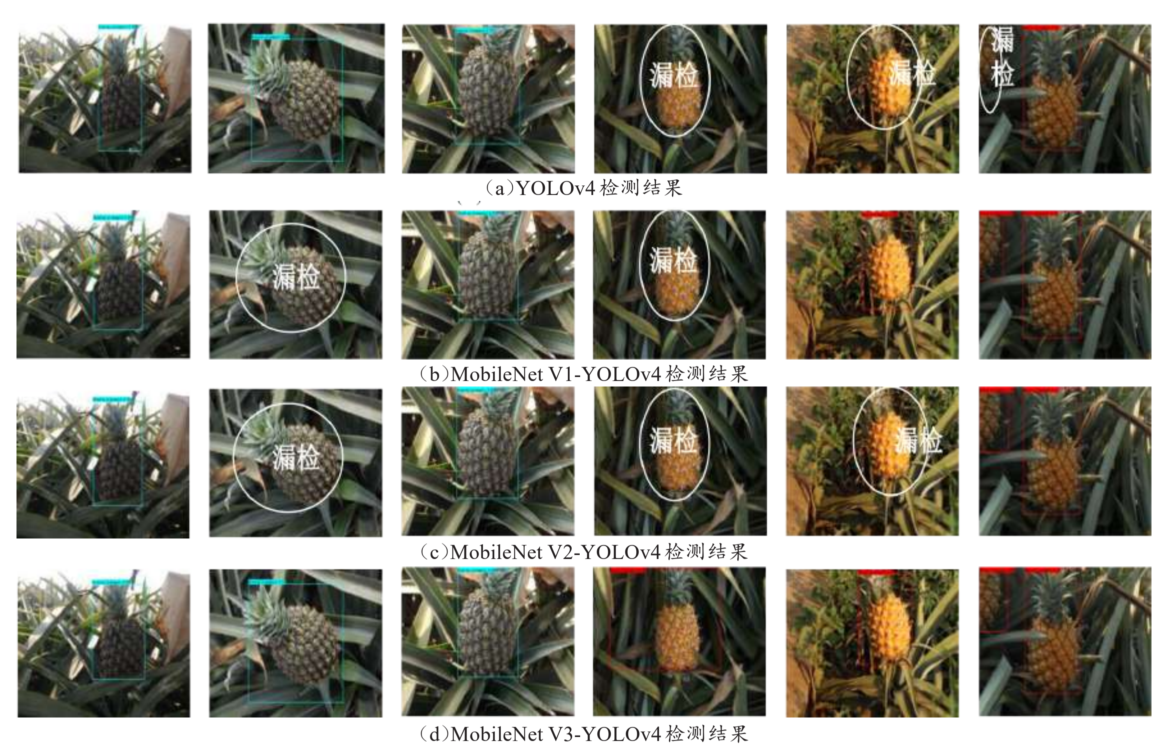
（4）从模型识别目标的速率考虑，将FPS作为评价指标。MobileNet V3-YOLOv4的FPS在YOLOv4系列表现最优，约为原YOLOv4网络推理速度的2倍，将改进后的模型应用到实际生产中，极大程度降低了时间成本。

综合以上所有因素，MobileNet V3-YOLOv4可以兼顾各种因素作为菠萝成熟度分析的可行模型。

3.4 复杂环境下MobileNet V3−YOLOv4检测菠萝结果

菠萝生长在环境复杂的农田场景中，影响菠萝的检测结果因素包括交叉的枝叶遮蔽、果实个体不完整、与自身颜色相似的背景颜色（主要针对青熟期菠萝）、其他的背景杂物，以及光照变化使得果实表面产生的阴影、光照较暗、光照较亮等。为了测试MobileNet V3-YOLOv4在复杂环境下对不同成熟度菠萝果实的检测效果，从本研究菠萝数据集的测试集中随机选取包含上述影响因素的照片进行检测，且每张照片中的影响因素并非单一存在，对应了真实的自然环境，验证改进后的模型在受到各种复杂因素影响下的自然环境中是否能取得较好的检测效果，漏检情况已用白色椭圆框标出，检测结果如图7所示。从测试集中挑选出6张图片进行检测，图片中共有7个菠萝，通过几种网络检测结果进行对比发现：（1）YOLOv4网络对于青熟期菠萝有较好的检测效果，但对于黄熟期菠萝检测效果较差，出现的漏检情况较其他三种网络最多，本研究的检测结果示例中有三处漏检的情况，漏检比例为3/7，与MobileNet V2-YOLOv4相同，在四种网络中最高；（2）MobileNet V1-YOLOv4和MobileNet V2-YOLOv4对于黄熟期的菠萝检测效果虽然有一定程

度的提升，但是也存在一定的漏检情况，对于黄熟期菠萝本研究的检测结果示例中MobileNet V1-YOLOv4有一处漏检的情况，MobileNet V2-YOLOv4有两处漏检的情况，且两种网络对于青熟期的菠萝也存在漏检情况，对于青熟期的菠萝，本研究的检测结果示例中两种网络均有一处漏检的情况；(3)与YOLOv4、MobileNet V1-YOLOv4和MobileNet V2-YOLOv4三种网络相比，MobileNet V3-YOLOv4可以兼顾两种成熟度的菠萝，表现效果最佳，符合自然环境下菠萝果实的检测需求。



注：左侧三列为青熟期菠萝，右侧三列为黄熟期菠萝

图7　四种网络在复杂农田场景下对不同成熟度菠萝的检测结果

Fig. 7 Detection results of pineapple with different maturity by four networks in complex farmland scenario

4　结　论

本研究通过将YOLOv4网络中的骨干网络替换为MobileNet V3，提出了轻量级的MobileNet V3-YOLOv4，利用复杂农田场景下拍摄的不同成熟程度菠萝果实照片进行训练测试，检验其性能优越性。通过对比发现MobileNet V3-YOLOv4网络的mAP值比YOLOv4网络高出5.28%，推理速度比原网络高出40.28 img/s，对青熟期的检测精确率比原网络高出3.34%，对黄熟期的菠萝检测精确率比YOLOv4高出6.25%，但其参数量只占YOLOv4网络的1/5，实现了优化模型精度和推理速度的同时减少了计算量和参数量，使得模型更易嵌入到不同的硬件平台，可以在复杂的自然场景下对菠萝的成熟度进行实时分析。

本研究的成熟度类型只有两种，在未来的研究中可以扩充数据集，加入不同的成熟度类型。本研究每张图像中均只有1～2颗菠萝果实，而实际的生产生活中，视野较广的农田只会架设少数几个监控摄像，监控摄像所拍摄的果实数量众多，图片放大后单一果实目标可能会像素化，从而影响网络的检测效果。通过提升硬件检测设备的性能，提升图片的分辨率，可实现图像放大后的精准检测。数据采集工作时为晴朗天气，实际的菠萝检测和采收作业可能发生在任何天气情况下，故天气因素也可作为数据集的影响因素之一，采集不同天气下图像会使得数据集更多样更丰富。未来还可以将MobileNet V3-YOLOv4网络与遥感技术、无人机等相结合，进一步实现整片农田菠萝成熟度实时分析，为菠萝实现自动化采摘提供一种理论支持。

利益冲突声明：本研究不存在研究者以及与公开

研究成果有关的利益冲突。

参考文献：

[1] 金琰, 刘海清, 刘恩平. 我国菠萝种植区域优势研究[J]. 中国农业资源与区划, 2014, 35(3): 100-104.
JIN Y, LIU H Q, LIU E P. Study on the regional advantages of pineapple planting in China[J]. Chinese journal of agricultural resources and regional planning, 2014, 35(3): 100-104.

[2] 刘传和, 刘岩, 谢盛良, 等. 不同成熟度菠萝果实香气成分分析[J]. 热带作物学报, 2009, 30(2): 234-237.
LIU C H, LIU Y, XIE S L, et al. Analysis of aroma components of pineapple fruit at different ripening stages[J]. Chinese journal of tropical crops, 2009, 30(2): 234-237.

[3] 刘胜辉, 孙伟生, 陆新华, 等. 6个菠萝品种成熟果实香气成分分析[J]. 热带作物学报, 2015, 36(6): 1179-1185.
LIU S H, SUN W S, LU X H, et al. Analysis of aromatic components in ripe fruits of 6 pineapple cultivars[J]. Chinese journal of tropical crops, 2015, 36(6): 1179-1185.

[4] FORTUNA-CERVANTES J M, RAMÍREZ-TORRESMT, MARTINEZ-CARRANZA J, et al. Object detection in aerial navigation using wavelet transform and convolutional neural networks: A first approach[J]. Proceedings of the institute for system programming of the RAS, 2021, 33 (2): 149-162.

[5] GAI R , CHEN N , YUAN H .A detection algorithm for cherry fruits based on the improved YOLOv4 model[J]. Neural computing and application, 2021, 35(19): 13895-13906.

[6] CHEN M C , CHENG Y T , LIU C Y .Implementation of a fruit quality classification application using an artificial intelligence algorithm[J]. Sensors and materials: An international journal on sensor technology, 2022, 34(1), 151-162.

[7] KUZNETSOVA A, MALEVA T, SOLOVIEV V. Using YOLOv3 algorithm with pre- and post-processing for apple detection in fruit-harvesting robot[J]. Agronomy, 2020, 10(7): 1016-1035.

[8] ZHANG Y D, DONG Z C, CHEN X Q, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation[J]. Multimedia tools and applications, 2019, 78(3): 3613-3632.

[9] CHAIKAEW A, THANAVANICH T, DUANGTANG P, et al. Convolutional neural network for pineapple ripeness classification machine[C]// 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). Piscataway, NJ, USA: IEEE, 2020: 373-376.

[10] CUONGN H H, TRINHTH, MEESAD P, et al. Improved YOLO object detection algorithm to detect ripe pineapple phase[J]. Journal of intelligent& fuzzy systems: Applications in engineering andtechnology, 2022, 43(1): 1365-1381.

[11] 张星, 高巧明, 潘栋, 等. 基于改进YOLOv3的田间复杂环境下菠萝拾捡识别研究[J]. 中国农机化学报, 2021, 42 (1): 201-206.
ZHANG X, GAO Q M, PAN D, et al. Picking recognition research of pineapple in complex field environment based on improved YOLOv3[J]. Journal of Chinese agricultural mechanization, 2021, 42(1): 201-206.

[12] LIU T H, NIE X N, WU J M, et al. Pineapple (*Ananas comosus*) fruit detection and localization in natural environment based on binocular stereo vision and improved YOLOv3 model[J]. Precision agriculture, 2023, 24(1): 139-160.

[13] 茅智慧, 朱佳利, 吴鑫, 等. 基于YOLO的自动驾驶目标检测研究综述[J]. 计算机工程与应用, 2022, 58(15): 68-77.
MAO Z H, ZHU J L, WU X, et al. Review of YOLO based target detection for autonomous driving[J]. Computer engineering and applications, 2022, 58(15): 68-77.

[14] WANG X, GAO J S, HOU B J, et al. A lightweight modified YOLOX network using coordinate attention mechanism for PCB surface defect detection[J]. IEEE sensors journal, 2022, 22(21): 20910-20920.

[15] XU D Q, WU Y Q. Improved YOLOv3 with DenseNet for multi-scale remote sensing target detection[J]. Sensors, 2020, 20(15): ID 4276.

[16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Conferenceon Computer Visionand Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.

[17] SUN Y Y, MA S, SUN S Y, et al. Partial discharge pattern recognition of transformers based on MobileNets convolutional neural network[J]. Applied sciences, 2021, 11(15): ID 6984.

[18] LI J P, ZHU K Y, WANG F, et al. Deep neural network-based real time fish detection method in the scene of marine fishing supervision[J]. Journal of intelligent& fuzzy systems: Applications in engineering and technology, 2021, 41(3): 4527-4532.

Pineapple Maturity Analysis in Natural Environment Based on MobileNet V3-YOLOv4

LI Yangde[1,2], MA Xiaohui[3*], WANG Ji[3,4*]

(1. *The Fourth Branch of Guangdong Telecom Planning and Design Institute Co. , LTD*, *Zhanjiang* 524088, *China*; 2. *Huazhong University of Science and Technology*, *Wuhan* 430070, *China*; 3. *College of Electronic and Information Engineering*,

Guangdong Ocean University, Zhanjiang 524088, *China*; 4. *Guangdong Intelligent Ocean Sensor Network and Equipment Engineering Technology Research Center, Zhanjiang* 524088, *China*）

Abstract:

[Objective] Pineapple is a common tropical fruit, and its ripeness has an important impact on the storage and marketing. It is particularly important to analyze the maturity of pineapple fruit before picking. Deep learning technology can be an effective method to achieve automatic recognition of pineapple maturity. To improve the accuracy and rate of automatic recognition of pineapple maturity, a new network model named MobileNet V3-YOLOv4 was proposed in this study.

[Methods] Firstly, pineapple maturity analysis data set was constructed. A total of 1580 images were obtained, with 1264 images selected as the training set, 158 images as the validation set, and 158 images as the test set. Pineapple photos were taken in natural environment. In order to ensure the diversity of the data set and improve the robustness and generalization of the network, pineapple photos were taken under the influence of different factors such as branches and leaves occlusion, uneven lighting, overlapping shadows, etc. and the location, weather and growing environment of the collection were different. Then, according to the maturity index of pineapple, the photos of pineapple with different maturity were marked, and the labels were divided into yellow ripeness and green ripeness. The annotated images were taken as data sets and input into the network for training. Aiming at the problems of the traditional YOLOv4 network, such as large number of parameters, complex network structure and slow reasoning speed, a more optimized lightweight MobileNet V3-YOLOv4 network model was proposed. The model utilizes the benck structure to replace the Resblock in the CSPDarknet backbone network of YOLOv4. Meanwhile, in order to verify the effectiveness of the MobileNet V3-YOLOv4 network, MobileNet V1-YOLOv4 model and MobileNet V2-YOLOv4 model were also trained. Five different single-stage and two-stage network models, including R-CNN, YOLOv3, SSD300, Retinanet and Centernet were compared with each evaluation index to analyze the performance superiority of MobileNet V3-YOLOv4 model.

[Results and Discussions] MobileNet V3-YOLOv4 was validated for its effectiveness in pineapple maturity detection through experiments comparing model performance, model classification prediction, and accuracy tests in complex pineapple detection environments.The experimental results show that, in terms of model performance comparison, the training time of MobileNet V3-YOLOv4 was 11,924 s, with an average training time of 39.75 s per round, the number of parameters was 53.7 MB, resulting in a 25.59% reduction in the saturation time compared to YOLOv4, and the parameter count accounted for only 22%. The mean average precision (mAP) of the trained MobileNet V3-YOLOv4 in the verification set was 53.7 MB. In order to validate the classification prediction performance of the MobileNet V3-YOLOv4 model, four metrics, including *Recall* score, F_1 Score, *Precision*, and average precision (AP), were utilized to classify and recognize pineapples of different maturities. The experimental results demonstrate that MobileNet V3-YOLOv4 exhibited significantly higher *Precision*, AP, and F_1 Score the other. For the semi-ripe stage, there was a 4.49% increase in AP, 0.07 improvement in F_1 Score, 1% increase in *Recall*, and 3.34% increase in *Precision* than YOLOv4. As for the ripe stage, there was a 6.06% increase in AP, 0.13 improvement in F_1 Score, 16.55% increase in *Recall*, and 6.25% increase in *Precision*. Due to the distinct color features of ripe pineapples and their easy differentiation from the background, the improved network achieved a precision rate of 100.00%. Additionally, the mAP and reasoning speed (Frames Per Second, FPS) of nine algorithms were examined. The results showed that MobileNet V3-YOLOv4 achieved an mAP of 90.92%, which was 5.28% higher than YOLOv4 and 3.67% higher than YOLOv3. The FPS was measured at 80.85 img/s, which was 40.28 img/s higher than YOLOv4 and 8.91 img/s higher than SSD300. The detection results of MobileNet V3-YOLOv4 for pineapples of different maturities in complex environments indicated a 100% success rate for both the semi-ripe and ripe stages, while YOLOv4, MobileNet V1-YOLOv4, and MobileNet V2-YOLOv4 exhibited varying degrees of missed detections.

[Conclusions] Based on the above experimental results, it can be concluded that MobileNet V3-YOLOv4 proposed in this study could not only reduce the training speed and parameter number number, but also improve the accuracy and reasoning speed of pineapple maturity recognition, so it has important application prospects in the field of smart orchard. At the same time, the pineapple photo data set collected in this research can also provide valuable data resources for the research and application of related fields.

Key words: pineapple maturity; backbone network; MobileNet V3-YOLOv4; Faster R-CNN; SSD300; Retinanet; Centernet; lightweight